Article

Chinese Science Bulletin

SPECIAL ISSUE
Adaptive Evolution and Conservation Ecology of Wild Animals

June 2013 Vol.58 No.18: 2148–2153
doi: 10.1007/s11434-013-5696-3

Characterization of MHC class II A genes in Hainan Eld's deer (*Cervus eldi hainanus*)

LIU HongYi, YANG HaiQiong & GE YunFa*

The Key Laboratory of Conservation Biology for Endangered Wildlife of the Ministry of Education and State Conservation Center for Gene Resources of Endangered Wildlife, College of Life Sciences, Zhejiang University, Hangzhou 310058, China

Received August 8, 2012; accepted October 29, 2012; published online May 7, 2013

The major histocompatibility complex (MHC) genes play pivotal roles in the immune system of vertebrates against antigens. They are also significant indicators of genetic structure, and are vital to species-level population viability analyses and disease risk assessments. In this study, two *DRA* and two *DQA* sequences were isolated from Hainan Eld's deer (*Cervus eldi hainanus*) using rapid amplification of cDNA ends (RACE) and single-strand conformation polymorphism-heteroduplex (SSCP-HD) analysis. Nucleotide sequence analysis revealed large differences between the two *DQA* sequences, especially in their exon 2 regions, but only minimal differences between the variants of the *DRA* gene. Comparison of the predicted amino acid sequences of the *Ceel*-MHC class II A variants with those from six other species revealed that these molecules share high homology among ruminants. A phylogenetic tree of four class II A sequences from Hainan Eld's deer and the other species placed the newly identified *DQA* and *DRA* genes on two distinct branches (100%-supportively), and further divided the two *DQA* sequences into 98%-supportive *DQA1* and 99%-supportive *DQA2* clusters, respectively. Therefore, this study identified monomorphic *Ceel-DQA1* and *Ceel-DQA2* genes, and one dimorphic *Ceel-DRA* gene from Hainan Eld's deer.

Hainan Eld's deer, MHC class II A genes, HURRAH, locus isolation

Citation: Liu H Y, Yang H Q, Ge Y F. Characterization of MHC class II A genes in Hainan Eld's deer (*Cervus eldi hainanus*). Chin Sci Bull, 2013, 58: 2148–2153, doi: 10.1007/s11434-013-5696-3

In recent decades, genetic factors have been recognized as critical components of the interacting forces that contribute to the extinction of endangered species [1–3]. Information on the genetic structure and adaptive evolution of a given species is extremely important for the conservation of endangered animals [4–6]. The major histocompatibility complex (MHC), one of the most varied gene families in the vertebrate genome, is an essential part of the adaptive immune system and has proven to be a critical marker in the field of molecular genetics [4]. MHC genes encode cell-surface glycoproteins that are responsible for presenting antigens derived from pathogens or parasites to T-lymphocytes [7]. MHC class II molecules are membrane-bound heterodimeric glycoproteins consisting of an α chain and a β chain, both of which are encoded by MHC genes. The MHC genes evolve rapidly under the selection pressure exerted by pathogens; in the class II genes, this is particularly notable in exon 2, which encodes the peptide-binding region (PBR) [8].

Hainan Eld's deer, *Cervus eldi hainanus*, is a subspecies of Eld's deer that is endemic only to Hainan Island of China [9,10]. Due to habitat destruction and hunting, this species has experienced a population decline that has put it on the verge of extinction; it is presently listed as a Class I species under state protection, and it is only found in reserves nowadays [11–13]. Although the population has gradually increased under legal protection, Hainan Eld's deer is still at risk for diseases and parasites, such as necrobacillosis and tick infection [14]. However, the immune potential of this species, which is closely related to the population's ability to rejuvenate, has not been studied in detail. The MHC genes, which are significant indicators of the genetic structure and disease risk of a species, have been successfully used to probe the genetic variation and immune potential of many other

*Corresponding author (email: lscgeyf@zju.edu.cn)

endangered wildlife species [15,16]. Thus, gaining a better understanding of the MHC genes in Hainan Eld's deer could facilitate the management of this deer population.

In the present study, in an effort to understand the genetic structure and immune potential of Hainan Eld's deer, we isolated *Ceel*-MHC class II A (alpha) loci using a newly modified HURRAH protocol. This protocol consists of two steps, "HUR" (isolation of MHC cDNA sequences) and "RAH" (isolation of MHC DNA sequences and identification of MHC genes)[17], and uses them to characterize MHC genes based on information obtained from both cDNA and genomic DNA.

1 Materials and methods

1.1 Sample collection and DNA preparation

Five tissue samples were collected from dead male deer at the Bangxi Nature Reserve (19°24′N, 108°56′E) in the fall of 2010. The samples were stored at −80°C. RNA was extracted from the liver of one individual (BX42) using TRIzol. The RNA integrity and concentration were checked, and good quality RNA was reverse transcribed to cDNA. Genomic DNA was extracted from all five samples, using a standard proteinase K/phenol-chloroform protocol [18].

1.2 Isolation of MHC class II A cDNA sequences

To obtain the complete cDNA sequences of MHC class II A genes from Hainan Eld's deer (individual BX42), we employed two sets of primers. The first set designed based on conserved mammalian MHC cDNA sequences [17] was adopted to amplify partial cDNA sequences in order to validate universality of primers. After that, we conducted rapid amplification of cDNA ends (RACE) using GeneRacer kit to retrieve the 3′- and 5′-UTR sequences. The second set of primers was designed from the obtained cDNA sequences, and was used to amplify exons 1–4 of the isolated MHC genes. The primers used for each step are presented in Table 1. All of the obtained MHC cDNA sequences were checked using single-strand conformation polymorphism-hetero- duplex (SSCP-HD) analysis, and these sequences were cloned and sequenced.

1.3 Isolation of MHC DNA sequences and identification of MHC genes

The structures of the MHC genes were highly similar but the intron sizes varied, so we designed the intron-amplifying primers to span from exon 1 to exon 3, and we used a two-step method in which the annealing temperature was a consistent 68°C, but the extension times varied from 3 to 6 min (Table 2). Specific primers were designed to amplify *DQA* genomic exon 2 sequences because the intron 1, 2 regions differed significantly between the two *DQA* sequences. Conversely, the *DRA* genomic exon 2 sequences were amplified using universal primers devised from conserved intron 1, 2 sequences. The number of genomic exon 2 sequences in the individuals with the simplest genetic structure reflected the number of MHC class II A loci in Hainan Eld's deer. We identified the genotype of each individual using the SSCP-HD technique. To identify the various

Table 1 Primers used to amplify the Hainan Eld's deer MHC cDNA sequences

Procedure	Primers	Sequence (5′→3′)	Region	Size (bp)	T_m (°C)
Amplification of partial cDNA	DQA-up (Qa-up)	CAGCCCAACACCCTCATCTG	Exon 3	212	55.6
	DQA-dn (Qa-dn)	TCCAGGCCCCAGTGCTCC			
	DRA-up (Ra-up)	GTGGATATGGVR AAGARGGAGAC	Exon 2–4	521	53.6
	DRA-dn(Ra-dn)	ATGATGCCCACCAGASCCAC			
	DRA-up (Ra-up′)[a)]	CHAYACYCCRAWCACCAATG	Exon 2–4	387	57.3
Amplification of the 3′UTR	DQAF	YGTGGACAACATYTTTCCYCCTGTGAT	Exon 3–3′UTR	728	72→50[b)]
	GeneRacer3′ne[a)]	CGCTACGTAACGGCATGACAGTG			
	DRAF	TCCTCCRGAAGTGACYCTGCTC	Exon 3–3′UTR	871	72→50[b)]
Amplification of the 5′UTR	DQAR	RGTCAACCCCARGGCACAGA	Exon 3–5′UTR	719	72→50[b)]
	GeneRacer5′ne[a)]	GGACACTGACATGGACTGAAGGAGTA			
	DRAR	GAGAACTTGTCAATGAAGCAGATGAGTGT	Exon 3–5′UTR	503	72→50[b)]
Amplification of exon 2 used for SSCP[c)]	DQAe2up	CCWCCATGATGWGCYCCAGT	Exon 2	328	72→50[b)]
	DQAe2dn	ATCACRGGAGACTTGGAAAACACA			
	DRAe2up	GAATCAKGGGCTATCAAAGAGGA	Exon 2	316	72→50[b)]
	DRAe2dn	GCTCTCCCAGTTCCACAGGCT			

a) Nested primers were used for amplification. b) We used a Touchdown PCR profile, in which the annealing temperature decreased progressively. c) The PCR products used for SSCP consisted of partial sequences from exons 1 and 3, and all of exon 2.

Table 2 Primers used for amplification of the Hainan Eld's deer MHC DNA sequences

Procedure	Primers	Sequence (5′→3′)	Region	Size (bp)	T_m (°C)
Amplification of intron 1, 2	DQA-1-UP	GATGAGCCCCAGTGGAGGTGAAG	Exon 1–3	Variable [b)]	68
	DQA-1-DN	GACTGAGTGTCCATTCCTTAG			
	DQA-2-UP	GGGGCCCTCACCCTGACC	Exon 1–3	Variable [b)]	68
	DQA-2-DN	GACTGCATGCCCATTCTTCAA			
	DRA-1-UP	CCTACAGGAATCATGGGCTATCAA	Exon 1–3	2851	68
	DRA-1-DN	TTGCGGAAAAGGTGGTCATC			
	DRA-2-UP	TGATCAGCCTACAGGAATCAG	Exon 1–3	Variable [b)]	68
	DRA-2-DN	ACAGGCTTGTCTGGGAGCAGG			
Amplification of exon 2	DQA-1-E2-UP	TTTTTTTTTGTTGTTGCTGT	Exon 2	365	55
	DQA-1-E2-DN	GAAGTGGTGGATCAATAGAGG			
	DQA-1-E2-UP′ [a)]	TTTCCACATTTCTTCTCACCC	Exon 2	318	58
	DQA-2-E2-UP	ATGTGCCAAAGTGAAGCCA	Exon 2	430	55
	DQA-2-E2-DN	AGAAGGGAATGAAGCCTGA			
	DQA-2-E2-UP′ [a)]	CCTGCTCCTCACCCTCACT	Exon 2	299	58
	DQA-2-E2-DN′ [a)]	GAGAAGTAGAATGGTGGACAC			
	DRA-E2-UP	CTCTCCATCTTCCTCTCCTGG	Exon 2	430	55
	DRA-E2-DN	GAATGTGGGGTTTGCCTATGA			
	DRA-E2-DN′ [a)]	AGGAGGACTGGAGCAGAGAGA	Exon 2	326	62

a) Nested primers were used for amplification. b) The size differences resulted from intra-locus differences in intron size and incomplete sequencing of some clones.

alleles, the fragments corresponding to each SSCP banding pattern were sequenced. In addition, the sequences from heterozygous animals were cloned and analyzed directly. The new alleles were identified in accordance with the regulations formulated by the nomenclature committee [19].

1.4 Data analysis

The Lasergene sequence analysis software was used for sequence editing and contig assembly [20]. Nucleotide and amino acid sequences were aligned using ClustalW [21]. A bootstrap consensus tree was constructed using the neighbor-joining method in MEGA 4.0 [22].

2 Results

2.1 Gene structure

Four MHC class II cDNA sequences, comprising exons 1–4, were obtained from individual BX42. The obtained sequences included two *DQA* sequences with overall lengths of 768 bp and two *DRA* sequences of 762 bp. The difference in nucleotide composition between the two *DQA* exon 2 sequences was 22%. In contrast, the *DRA* exon 2 sequences differed by only 0.8%. The predicted amino acid sequences were analogous to related sequences from other ruminants (Figure 1). A GenBank search showed that the closest homologs to the Hainan Eld's deer sequences belonged to buffalo, sheep, cattle and goat, whose MHC class II A sequences clustered with those of Hainan Eld's deer in a phylogenetic analysis (Figure 2).

2.2 Loci and alleles

The phylogenetic tree showed that the newly identified four sequences were grouped into three high-supportive clusters (*DQA1*, *DQA2* and *DRA*) together with the A genes of other ruminants, showing 98%, 99% and 100%, bootstrap values, respectively (Figure 2). The minimum number of genomic exon 2 sequences identified from an individual was three as well, indicating that there are three MHC class II A genes in the Hainan Eld's deer genome. Therefore, we designated these novel A genes as *Ceel-DQA1*, *-DQA2* and *-DRA*, respectively. According to the linkage disequilibrium theory, we thus found two common MHC class II A haplotypes in the five samples, as verified by SSCP-HD and sequence analysis: *DQA1*(a)-*DQA2*(a)-*DRA*(a) and *DQA1*(a)-*DQA2*(a)-*DRA*(b). The individuals with the maximum number of sequences only had four, and identical alleles were found in most individuals; overall, the number of alleles was very limited and there were few heterozygotes (Figure 3).

3 Discussion

3.1 Characterization of the Hainan Eld's deer MHC II A genes

In this study, we characterized the MHC class II A genes in the endangered Hainan Eld's deer. The following three lines

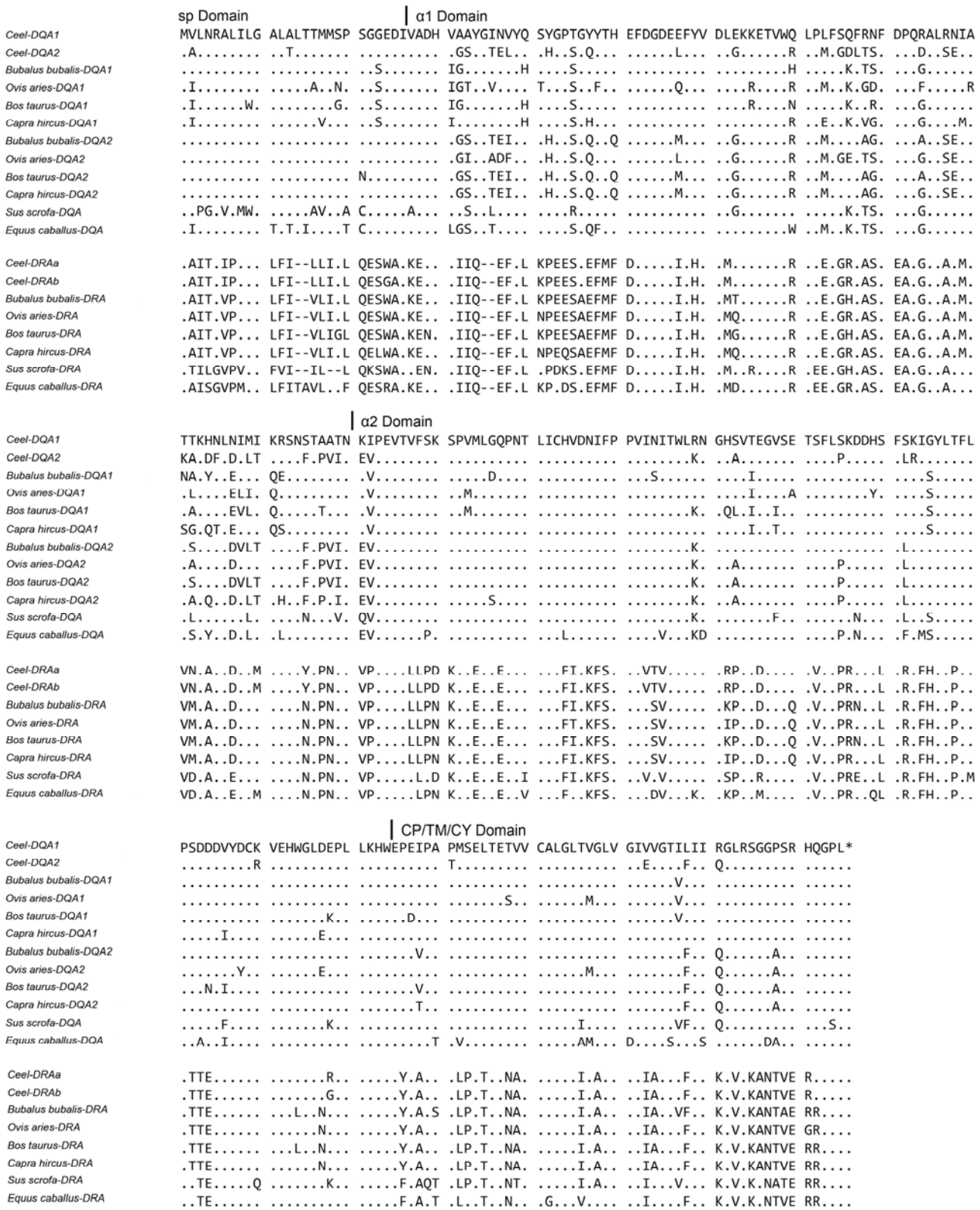

Figure 1 Alignment of predicted amino acid sequences for MHC class II A genes from Hainan Eld's deer and other vertebrates (*Bubalus bubalis*: DQ440647.1, DQ440648.1, DQ016629.1; *Ovis aries*: NM_001159759.1, M93431.1, JF735994.1; *Bos taurus:* BC146039.1, BC102953.1, NM_001012677.1; *Capra hircus*: AY464652.1, EU559625.1, AB008754.1; *Sus scrofa*: DQ003300.1, AY247782.1 *Equus caballus*: NM_001142814.1, XM_001494553.1).

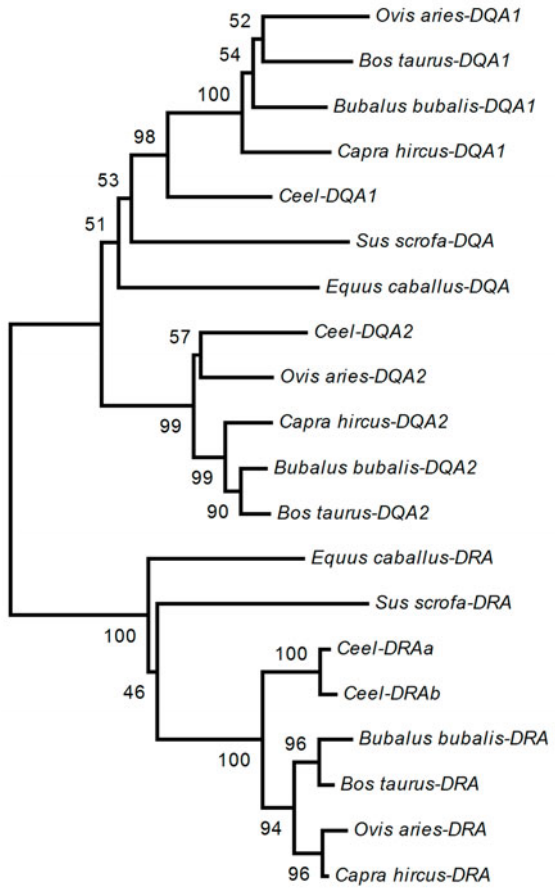

Figure 2 Phylogenetic tree of MHC class II A exon 1–4 nucleotide sequences from Hainan Eld's deer and other vertebrates. GenBank Accession numbers for the comparison sequences: *Bubalus bubalis-DQA1*, DQ440647.1; *Bubalus bubalis-DQA2*, DQ440648.1; *Bubalus bubalis-DRA*, DQ016629.1; *Ovis aries-DQA1*, NM_001159759.1; *Ovis aries-DQA2*, M93431.1; *Ovis aries-DRA*, JF735994.1; *Bos taurus-DQA1*, BC146039.1; *Bos taurus-DQA2*, BC102953.1; *Bos taurus-DRA*, NM_001012677.1; *Capra hircus-DQA1*, AY464652.1; *Capra hircus-DQA2*, EU559625.1; *Capra hircus-DRA*, AB008754.1; *Sus scrofa-DQA*, DQ003300.1; *Sus scrofa-DRA*, AY247782.1; *Equus caballus-DQA*, NM_001142814.1; *Equus caballus-DRA*, XM_001494553.1.

of evidence indicate that there are two *DQA* genes and one *DRA* gene in the genome of this species: (1) the nucleotide sequences of the two *DQA* exon 2 sequences differed significantly (by 22%) but those of two *DRA* exon 2 variants differed by only 0.8%; (2) our phylogenetic tree of theMHC class II A exon 1–4 nucleotide sequences separated the two *DQA* sequences and clustered the *DRA* sequences; (3) our SSCP-HD analysis showed that the simplest pattern only had three sequences in a single individual, and the five samples harbored two common haplotypes: *DQA1*(a)-*DQA2*(a)-*DRA*(a) and *DQA1*(a)-*DQA2*(a)-*DRA*(b). In addition, no additional alleles were found at DQA loci using the specific primers. Most small populations that have undergone bottlenecks (i.e. endangered animals) often show monomorphism at particular loci [17,23]. However, we predict that new alleles might be found in follow-up experiments with more samples of Hainan Eld's deer.

Our amino acid sequence and phylogenetic analyses revealed that the MHC class II A molecules shared high homologies among ruminants. There were clear differences between the Hainan Eld's deer sequences and those of the horse, whereas the deer sequences were very similar to those from the other ruminants. The phylogenetic tree of MHC class II A exon 1–4 nucleotide sequences separately clustered the *DQA* and *DRA* genes from ruminants, while the genes from the boar and the horse were positioned outside these branches. All of our experimental results were consistent with the taxonomy. We speculate that these findings may reflect that the tested ruminants have faced the same or similar pathogens and parasites in nature, leading to selection pressures that converge their MHC sequences.

3.2 Technology of MHC locus isolation

Previous papers have reported the isolation of MHC loci. Many of them used universal primers to amplify sequences from a group of individuals, then based their results on the maximum number of MHC sequences observed in any one individual. More recently, Wan et al. [17] introduced the new HURRAH technique for clearly characterizing MHC genes according to nucleotide sequences and the genotype of the individuals with the simplest pattern(s). This protocol was successfully employed to identify the MHC class II genes in both carnivore giant panda and herbivore David's deer [17]. Nevertheless, this HURRAH protocol uses a relatively complex magnetic bead hybridization to obtain full-length cDNA sequences since Wan et al. [17] aimed at developing a method able to isolate all potential sequences from a species with no or few homologous sequences. In this study, in view of cDNA sequences available from cervids, we used RACE instead of magnetic bead hybridization, thereby simplifying the isolation of cDNA sequences. As in the previous HURRAH protocol, we used SSCP-HD for genotyping and locus isolation. The most convincing evidence for the validity of our MHC locus identification came from the genotypes of the individuals with the simplest SSCP-HD banding patterns. For instance, we designed universal primers based on the conserved regions of the two *DRA* sequences, used them to amplify genomic sequences for *DRA* exon 2, and subjected the amplified fragments to SSCP-HD separation. The SSCP-HD banding patterns showed that two of the five individuals were homozygous, which meant that the two *DRA* sequences isolated from individual BX 42 were alleles, and there was only one *DRA* gene in the Hainan Eld's deer genome. Thus, our experimental results indicate that the SSCP-HD technique is a reliable method for genotyping, and that our newly-modified HURRAH protocol is appropriate for the isolation of MHC loci.

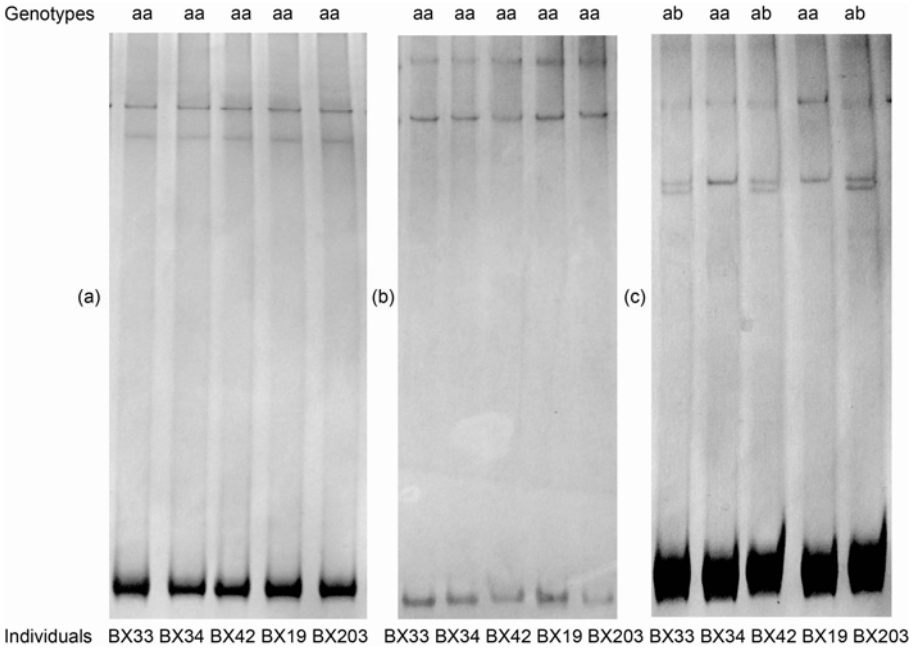

Figure 3 SSCP banding patterns of genomic exon 2 PCR products from Hainan Eld's deer. (a) *Ceel-DQA1*; (b) *Ceel-DQA2*; (c) *Ceel-DRA*.

We are grateful to the Bangxi Nature Reserve for providing us with the precious samples. This work was supported by the National Natural Science Foundation of China (30970426), a special grant from the State Forestry Administration of China, and the Fundamental Research Funds for the Central Universities of China.

1 Booy G, Hendriks R J J, Smulders M J M, et al. Genetic diversity and the survival of populations. Plant Biol, 2000, 2: 379–395
2 Reed D H, Frankham R. Correlation between fitness and genetic diversity. Conserv Biol, 2003, 17: 230–237
3 Spielman D, Brook B W, Frankham R, et al. Most species are not driven to extinction before genetic factors impact them. Proc Natl Acad Sci USA, 2004, 101: 15261–15264
4 Sommer S. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Front Zool, 2005, 2: 16
5 Lacy R C. Loss of genetic diversity from managed populations: Interacting effects of drift, mutation, immigration, selection, and population subdivision. Conserv Biol, 1987, 1: 143–158
6 Pan H J, Shi F L, Chang Z F, et al. Mitochondrial DNA variation analysis suggests extreme low genetic diversity in Guizhou snub-nosed monkeys (*Rhinopithecus brelichi*). Chin Sci Bull, 2011, 56: 2541–2544
7 Klein J. Natural History of the Major Histocompatibility Complex. New York: Wiley, 1986
8 Reche P A, Reinherz E L. Sequence variability analysis of human class I and class II MHC molecules: Functional and structural correlates of amino acid polymorphisms. J Mol Biol, 2003, 331: 623–641
9 Ohtaishi N, Gao Y T. A review of the distribution of all species of deer (Tragulidae, Moschidae and Cervidae) in China. Mamm Rev, 1990, 20: 125–144
10 Balakrishnan C N, Monfort S L, Gaur A, et al. Phylogeography and conservation genetics of Eld's deer (*Cervus eldi*). Mol Ecol, 2003, 12: 1–10
11 Song Y L. Population viability analysis for two isolated populations of Haianan Eld's deer. Conserv Biol, 1996, 10: 1467–1472
12 Zeng Z G, Song Y L, Li J S, et al. Distribution, status and conservation of Hainan Eld's deer (*Cervus eldi hainanus*) in China. Folia Zool, 2005, 54: 249–257
13 Pang J F, Hoelzel A R, Song Y L, et al. Lack of mtDNA control region variation in Hainan Eld's deer: Consequence of a recent population bottleneck? Conserv Genet, 2003, 4: 109–112
14 Lin X M. Incidence of diseases of Eld's deer (*Cervus eldi hainanus*) and strategies of prevention (in Chinese). Chin J Wildl, 2008, 29: 174–176
15 Ellegren H, Hartman G, Johansson M, et al. Major histocompatibility complex monomorphism and low levels of DNA fingerprinting variability in a reintroduced and rapidly expanding population of beavers. Proc Natl Acad Sci USA, 1993, 90: 8150–8153
16 Hedrick P W, Lee R N, Parker K M, et al. Major histocompatibility complex (MHC) variation in the endangered Mexican wolf and related canids. Heredity, 2000, 85: 617–624
17 Wan Q H, Zhang P, Ni X W, et al. A novel HURRAH protocol reveals high numbers of monomorphic MHC class II loci and two asymmetric multi-locus haplotypes in the Père David's deer. PLoS One, 2011, 6: e14518
18 Sambrook J, Russell D W. Molecular Cloning: A Laboratory Manual. New York: CSHL Press, 1989
19 Kennedy L J, Altet L, Angles J M, et al. Nomenclature for factors of the dog major histocompatibility system (DLA), 1998. First report of the ISAG DLA Nomenclature Committee. Tissue Antigens, 1999, 54: 312–321
20 Burland T G. DNAStar's Lasergene sequence analysis software. Methods Mol Biol, 2000, 132: 71–91
21 Aiyar A. The use of Clustal W and Clustal X for multiple sequence alignment. Methods Mol Biol, 2000, 132: 221–241
22 Tamura K, Dudley J, Nei M, et al. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol, 2007, 24: 1596–1599
23 Chen Y Y, Zhang Y Y, Zhang H M, et al. Natural selection coupled with intragenic recombination shapes diversity patterns in the major histocompatibility complex class II genes of the giant panda. J Exp Zool B Mol Dev Evol, 2010, 314: 208–223